Chapter 4: How the developing song is compared to the song model

4.1 Background and Rational

In Chapter 3 we discussed the difficulties of quantifying the variability across renditions of syllables. We introduced the method for identifying and tracking of the intra-syllabic events which ultimately can be used to compute the unbiased estimates of exploratory variability on short time scales (short lasting events inside of a syllable).

In the present chapter we turn to our main question: do birds regulate exploratory variability on a short time scale, within a syllable (which is a true continuous action). As discussed in the Introduction birds posses the brain pathway responsible for song learning (AFP), which can induce fine time-scale changes to the song performance (Kao et al., 2005) and this can be used for exploration of acoustic space. So while the “hardware” seems to be in place for such localized application of exploratory variability we do not know if birds actually use it this way.

Another issue that we discuss in this chapter is the question of diurnal oscillations of syllabic structure and the role of exploratory variability in the morning singing (when variability is higher).

With our method for measuring the exploratory variability of intra-syllabic events at hand we now look at how variability correlates with acoustic error. But first let us define “local error” and “global error”.

4.2 Methods and definitions

4.2.1 Local acoustic error of intra-syllabic events

Local error was obtained by computing the Euclidean distances for each intra-syllabic event relative to the corresponding events in the target. Distances were computed for each day using the daily means of acoustic features. These daily measures provide a time course of the local errors over development. This time course was correlated with the variability of intra-syllabic events which we call the *local correlation* (r_{local}).

4.2.2 Global acoustic error of intra-syllabic events

To estimate the global errors for intra-syllabic events, we computed the mean for each syllable of the local acoustic errors across intra-syllabic events. Again, these values are obtained for each day and correlated with the variability of intra-syllabic events which we call now the *global correlation* (r_{global}). Notice that we obtained these global correlations for each intra-syllabic event separately.

4.3 Results

4.3.1 Correlating exploratory variability with acoustic error

To visualize an example of the detailed relationship between variability and distance from target we plotted them against each other as shown for one syllable in Figure 4.1*A*, *B*. Distances of Wiener entropy and mean frequency were computed in reference to the end-point (Fig. 4.1*A*) and also in reference to the song model (Fig. 4.1*B*). As shown, in all intra-syllabic events, variability decreased across the time of development, but at different rates. For most events there was an association between the distance from the song model (or from the end-point) and the level of variability. For example, looking at Fig. 4.1*B*, for events 1 and 6, the initial variability was high in both, but in event 1 variability and distance from target decreased substantially in approximately five days, whereas in event 6 variability and distances both remained at high levels for about 20 days. Another interesting case is event 3, which bifurcated after 10 days into events 4 and 5. Note that event 4 was initially closer to the target and also of lower variability than event 5.

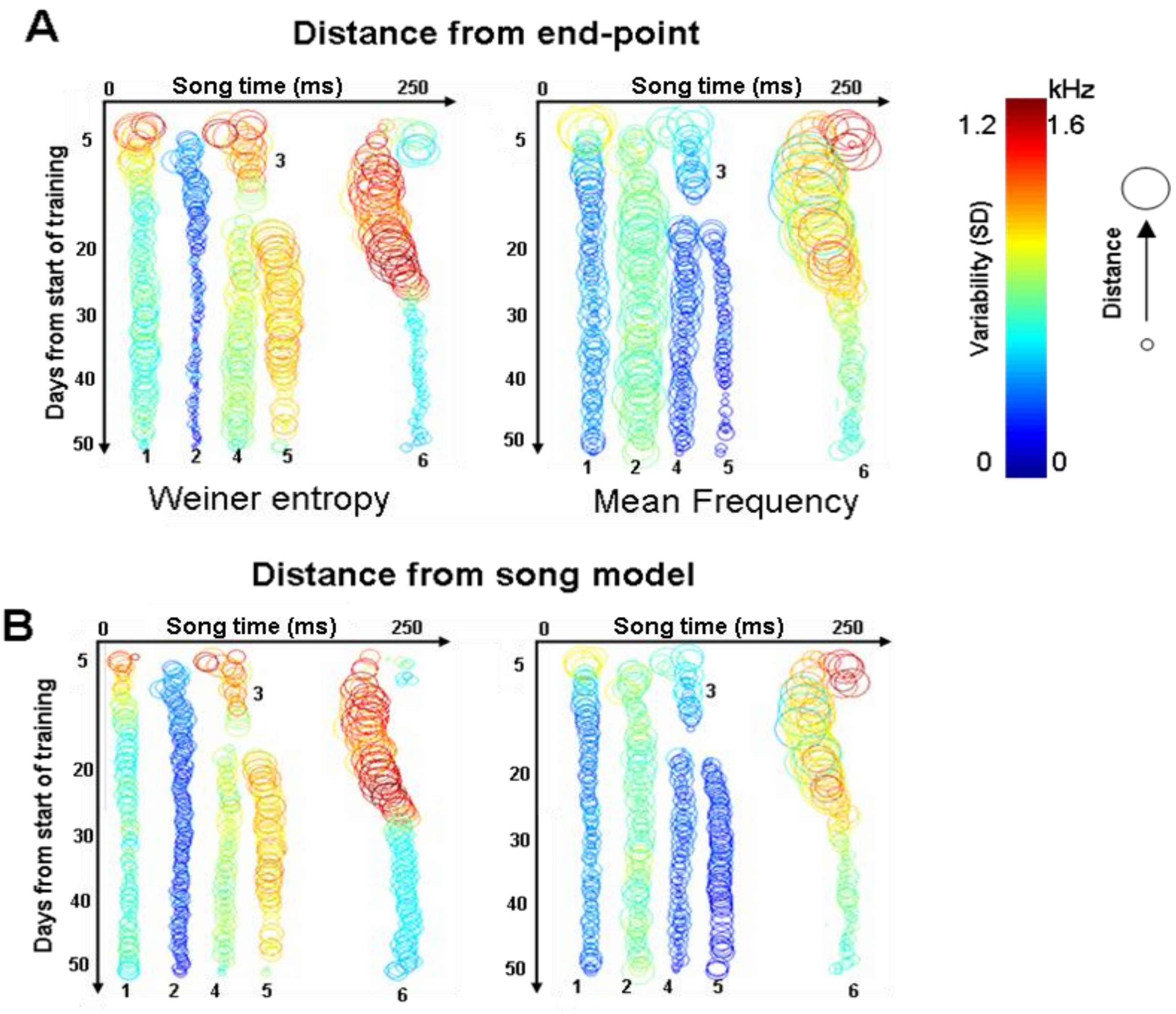

Figure 4.1 Variability decreases when intra-syllabic events reach their targets. ***A,*** Across development, variability of significant events decreases near their local end-points. Color of circles corresponds to $SD_{intrasyll}$ and size represents the Euclidian distance of intra-syllabic events from their end-points. ***B,*** The same variability data as in ***A*** is presented with the Euclidian distance from the corresponding intra-syllabic event in the song model.

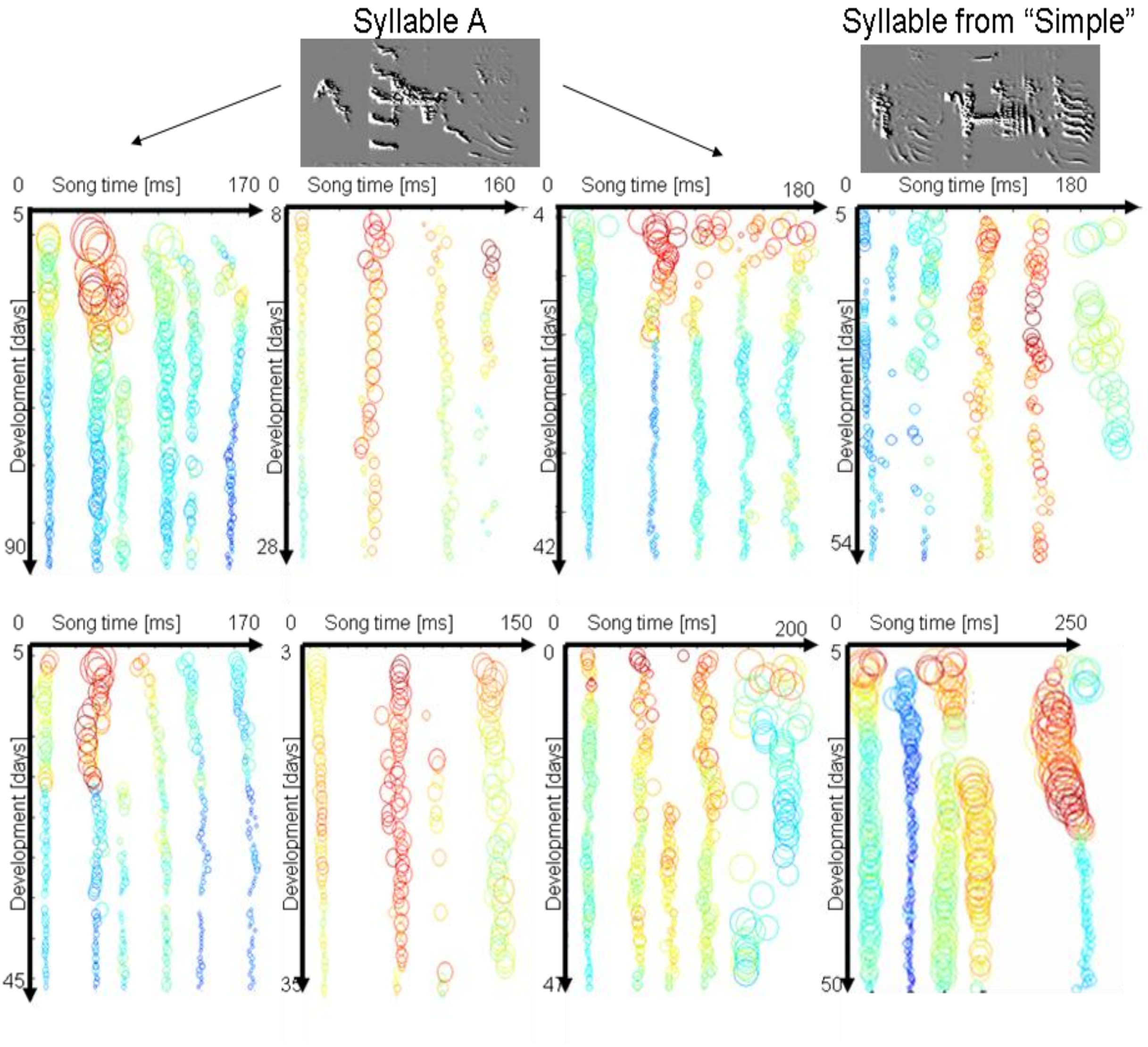

Figure 4.2 Exploratory variability and distance from target across birds. In most intra-syllabic events, across all eight birds studied the magnitude of variability $SD_{intrasyll}$ and the distance from the target are correlated. As in Fig 4.1 color of circles corresponds to $SD_{intrasyll}$ and size represents the Euclidian distance of intra-syllabic events from their corresponding events in the target song. The feature shown is Wiener entropy. Six birds were trained with sequential training (see section 2.2.4) and two birds were trained with a single song model (“Simple”). For both models the most complex syllables was used in the analysis.

Figure 4.2 shows the how exploratory variability relates to the distance from the target across all eight birds used in this study.

4.3.2 Variability correlates better with local error than with global error

We wondered if the magnitude of vocal exploration differs across intra-syllabic events and if so, whether different time courses of variability might mirror differences in the learning pace of different intra-syllabic events, i.e. whether decrease in acoustic error drives adaptive decrease in variability separately for each event. To test for this we computed the Euclidean distance of intra-syllabic events from the corresponding events in the target (*local error*). We then examined if this local error can explain the local variability ($SD_{intrasyll}$) better than the *global error* over the entire syllable. Note that the global error (being the average error of intra-syllabic events in the syllable) provides a more stable estimate compared to individual local errors, and therefore by default should provide slightly better correlations. We first computed the time course of global error across development, and correlated it with the time course of variability for each intra-syllabic event ($SD_{intrasyll}$). This estimate, r_{global}, was then compared to the correlation between the local errors and local variability, r_{local}.

To test if exploratory variability correlates locally within a syllable, we analyzed eight complex syllables, one from each bird trained with either sequential training procedure (six birds) or a single song model (two birds, see Method sections in Chapter 3, 3.2.1 & 3.2.2). The eight syllables selected (always the most complex syllable in a song in each bird) typically had 4-6 intra-syllabic events. The relationships between distance from the target and variability ($SD_{intrasyll}$) intra-syllabic events, across all eight birds are shown in Figure 4.2.

We computed r_{local} and r_{global} for each intra-syllabic event in each of the eight syllables, which produced 41 values for r_{local} and r_{global} .

Figure 4.3 shows plots of r_{local} vs. r_{global}, for Wiener entropy, mean frequency and time position. The data include all identified intra-syllabic events across all birds (n=41 intra-syllabic events from 8 birds). Measures of intra-syllabic events within a syllable cannot be considered as independent samples. Therefore, for the purpose of statistical testing, we computed median correlations (r_{local} and r_{global}) across intra-syllabic events for each bird, to obtain a single statistical estimate per bird. We restricted this analysis to differences from the song model (differences from the endpoint were only inspected qualitatively). <u>Variability was significantly more correlated with local error than with global error in all song features tested</u>: Wiener entropy (p= 0.015) mean frequency (p= 0.04) and time positions (p= 0.044, paired, two-tailed *t-test*; n=8 birds). Results in reference to end-points and song model were similar, but there were more cases of negative correlations when the reference was a song model. Investigating these showed that the bird had reached the model values, but then continued with vocal changes further away from the model, perhaps purposely diverging from the model.

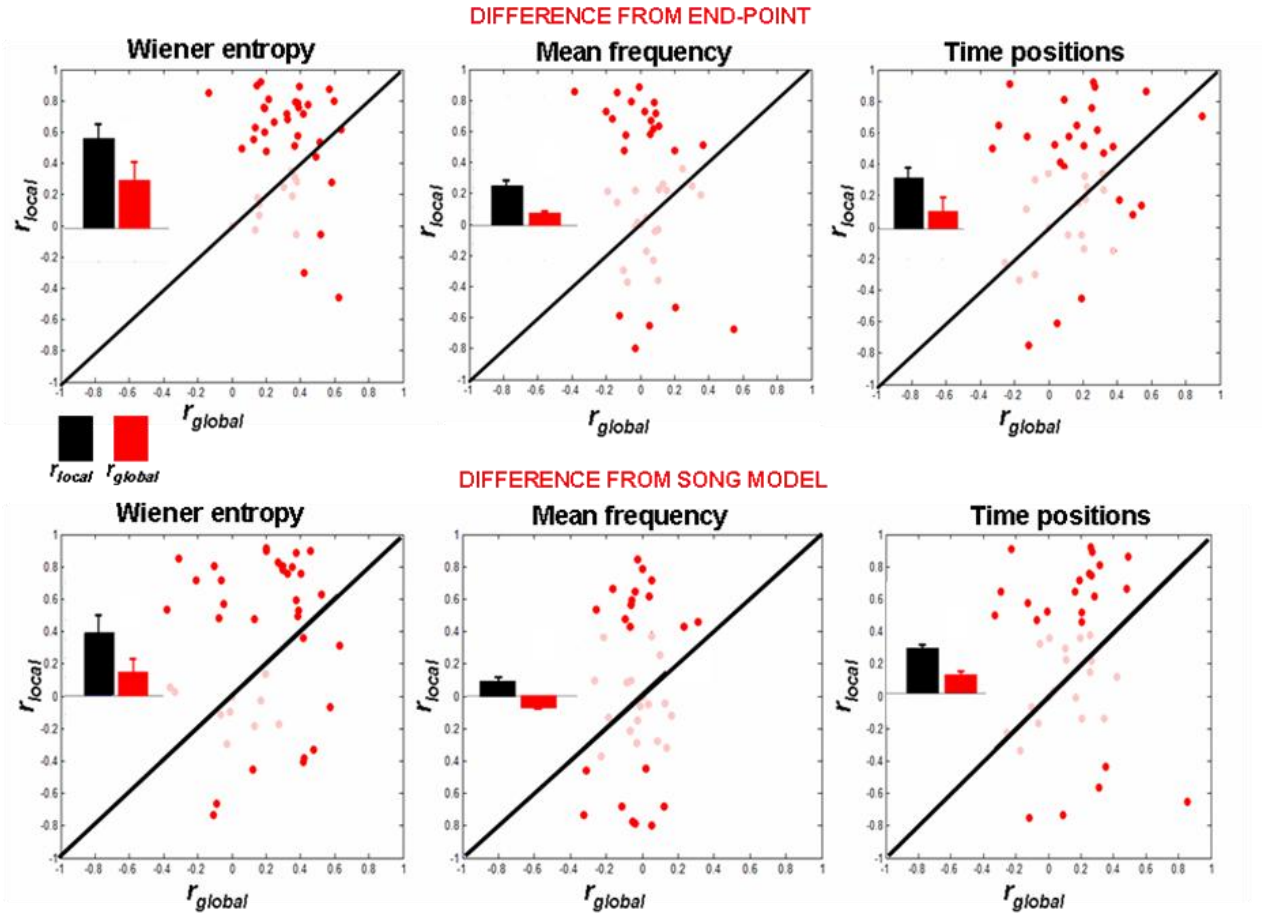

Figure 4.3 Variability of intra-syllabic events is better correlated with local than with global error. We correlated time courses of variability ($SD_{intrasyll}$) of intra-syllabic events with their local error to obtain r_{local} and with global error to obtain r_{global}. The errors were computed for Euclidian distances from both end-point (top) and the song model (bottom). Local and global correlates (r_{local} and r_{global}) were computed for all intra-syllabic events, across 8 birds, separately for each feature. Local correlates are significantly greater than global correlates for all three song features (see text). For visualization, we denoted non-significant correlations by pink circles. The bars present mean and s.e.m for correlations with local (black) and global (red) correlations.

4.3.3 Exploratory variability changes simultaneously with local error

We next examined if changes in variability ($SD_{intrasyll}$) and local error occurred *simultaneously* within the one-day time resolution of our study. To test for this, we computed correlations introducing lags between error and variability (cross-correlations). We computed lags for all three song features in all intra-syllabic events (across all eight birds). We then computed histograms of lags for all three features (y-axis represents the number of events and x axis the lag). Thus we obtained three histograms, one for each song feature. Next we combined the three histograms (added the frequencies) and obtained a single histogram representing all three song features, as shown in Figure 4.4. There was no significant deviation of correlations from lag=0 (*p*=0.512, *paired t-test*). This result indicates that changes of variability and local error are indeed simultaneous within the one-day time resolution analyzed.

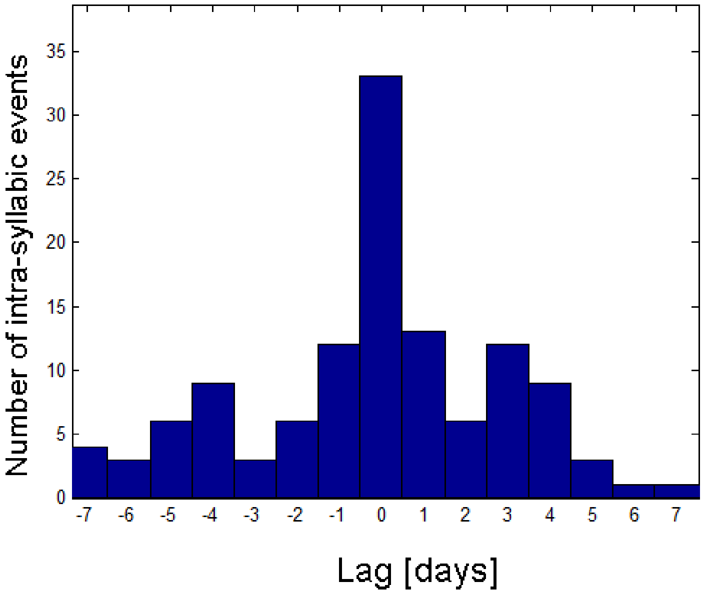

Figure 4. 4 Exploratory variability changes simultaneously with local error. We computed cross-correlations of local error and exploratory variability ($SD_{intrasyll}$) for all intra-syllabic events across all eight birds studied. Histograms of lags were computed for each of the three song features. In every feature the highest number of events had lag of zero. The histogram presented here is the cumulative histogram for all three features (sum of the three histograms).

An alternative explanation to our findings would be that perhaps certain song elements are both easier to learn and faster to stabilize (regardless of learning). To test for this we compared the learning of four equivalent intra-syllabic events in six birds that were trained with the same target syllable. We assessed the learning speed of each intra-syllabic event by computing the overall change in distance from the song model, and calculating the developmental time when the event reached half of this distance; *ANOVA* of these learned speed estimates showed no significant difference between the intra-syllabic events (*p*=0.124; *F*=2.25; across *n*=6 birds). Also note in Figure 4.2 that there seems to be obvious order at which intra-syllabic events consolidate in the six birds trained with syllable A (the sizes of the circles represent the distance from the target). Therefore, different birds learned similar intra-syllabic events in different relative rates, and the only correlation that holds is the one between the local error and variability.

Taken together these results indicate that the changes in local variability are best explained by the local error from the model, suggesting that birds can evaluate the error locally, at short time scales of no more than 20-50 ms, and maintain high exploratory variability primarily in those intra-syllabic events where local error is high. It is particularly interesting that variability not only of spectral features but also of time positions was better correlated with local error (of timing) than with global timing error, which suggests that time-jitter of intra-syllabic events is locally gated within a developing syllable.

4.3.4 Diurnal oscillations in variability

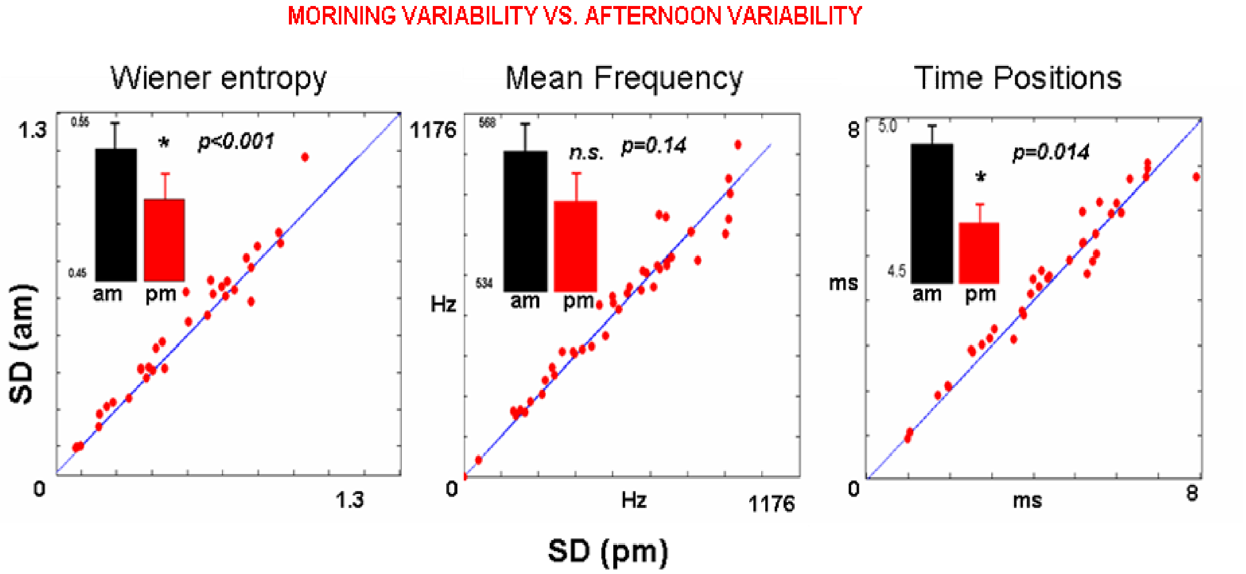

Figure 4. 5 Variability of intra-syllabic events tends to be higher in the morning than later in the day. Time courses of variability ($SD_{intrasyll}$) were computed for each intra-syllabic event from samples taken in the morning and was compared to total variability time course taken from late afternoon samples. In two song features (Wiener entropy and time positions) variability was significantly higher in the morning than in the afternoon in ($p<0.001$ for Wiener entropy and $p=0.02$ for time positions). For mean frequency the difference between morning and afternoon variability was not significant ($p=0.14$).

In the analyses thus far, with daily units taken from the afternoon and evening songs, developmental time courses of error and variability were often monotonic. However, previous investigations have shown that during periods of rapid learning, there are strong diurnal oscillations in song structure, such that the morning song is less structured and less similar to the song model (Derégnaucourt et al., 2005). We would predict that if a bird can adaptively change the magnitude of variability, the morning song, which is less similar to the model, should be more variable than the afternoon song. But a previous study (Miller et al., 2010) showed that syllables become more, rather than less, variable after daily practice (comparing morning to afternoon songs).

We therefore performed a similar analysis, but instead of examining variability at the syllable level, we computed variability of intra-syllabic events.

In order to assess diurnal effect we repeated the computation of variability as above, except that we segregated the data into the first and last thirds of the day (corresponding approximately to morning versus evening renditions). Note that estimates of variability obtained by our method do not depend on the alignment of syllables nor do they depend on the modulation of syllabic features. Highly structured syllables tend to be more modulated while in less structured syllables acoustic features are more “flat,” which could result in biased estimates of variability because higher diversity of vocal states (typical of highly developed syllables) may produce higher estimates of differences between syllabic renditions, which in turn may result in higher estimates of variability (across renditions). Identifying intra-syllabic events circumvents this problem.

As shown in Figure 4.5, variability across all intra-syllabic events is significantly higher in the mornings for Wiener entropy and time-positions, which is consistent with the hypothesis that young birds can increase or decrease variability adaptively.

4.4 Conclusion and discussion

Putting problems with articulatory dynamics (such co-articulation) aside, partitioning the continuous action could reduce the complexity of the song learning task.

Partitioning the learning of a continuous action into discrete tasks is sometimes considered suboptimal (Doya, 2000). At one extreme, the action can be segmented very coarsely and error can then be computed for each segment. At the extreme coarseness of segmentation (single segment) this would mean computing a global error. At the other extreme, the action can be segmented into a very high number of partitions, which would necessitate the computation of many local errors. This would require large memory capacity and many learning trials (Doya, 2000).

In zebra finch song, however, the overall number of distinct song elements (syllables and notes) ranges only from 5 to 20 and so the scale of partitioning does not need to be too fine. To illustrate the advantage of partitioning consider, for example, the learning of 10 syllables, each with 10 possible vocal events (such as our intra-syllabic events). Assuming that the bird learns these vocal events one by one, and notes do not interfere with each other (no co-articulation), then the bird needs to learn 10 tasks of 10 states (100 possible states). However, if the error is only available globally, than the bird would have to select among 10^{10} possible states (regardless of the precision of the error computation).

4.4.1 Age dependent plasticity and the gain of exploratory variability

Zebra finches rapidly lose their song learning abilities with age and with learning as their song turns from variable to stereotyped (Morrison and Nottebohm, 1993; Boettiger and Doupe, 2001), but they nevertheless retain some plasticity into adulthood. The minor residual variability in song features across renditions, that persists even in “stereotyped” song syllables, is still accessible to reinforcement learning, and can be used to train birds to shift the fundamental frequency of a targeted vocal element within syllable up or down (Tumer and Brainard, 2007; Andalman and Fee, 2009). In these experiments a short burst of white noise was played to a singing bird as a negative reinforcement, at the precise time when a particular vocal event was detected in the bird’s song. The negative reinforcement was removed if the bird either increased or decreased the fundamental frequency (to reach the “escape frequency” defined by the experimentalists). This training procedure resulted in consistent modification of the targeted vocal element. Interestingly, however, once the negative reinforcement training was finished, the fundamental frequency of the vocal element started to shift back to the original value. This suggests that the error (deviation from the original song model) is persistently reported after the song is modified. It seems that the memorized “template” does not change even if the song is modifiable in adult birds.

However, even prolonged training of adult birds which succeeded to shift their fundamental frequency did not induce any increments in the gain of variability in the targeted syllables (Tumer and Brainard, 2007). If the “escape frequency” that a bird had to reach (in the targeted vocal element) was too far from the original fundamental frequency, the animal never succeeded to modify the song and escape the negative reinforcement. This result proves that residual exploratory variability is indeed required for birds in order to locally modify their song but also implies that they can only do this in <u>small steps</u> (not exceeding the gain of residual variability). This also suggests that the developmental transition from high to low variability cannot be easily reversed. (For if it was easily reversible, then why would birds “choose” to not increase the exploratory variability of the targeted vocal element?)

These findings are reminiscent of the results by Knudsen and Knudsen, which show the differences in plasticity between young and old animals in terms of adaptive adjustment of auditory orienting behavior in response to displacing-prisms in barn owl (Knudsen and Knudsen, 1989). Juvenile owls can calibrate their auditory map to adjust for large angular errors in their visual field, but adult owls can only adapt to small errors and thus only learn in small steps. In the adult owl, the ability to adaptively rotate the auditory map is constrained by a narrowing integration window at the neuronal level. Analogously, in songbirds, as we have seen, it is the range of active vocal exploratory variability that constrains learning (Tumer and Brainard, 2007). Vocal exploration in juvenile birds provides a broader range of usable song elements than the adult song, due to stronger variability within a syllable and across syllabic renditions as well as stronger diurnal oscillation in song structure (Derégnaucourt et al., 2005).

Our interpretation of this result is that different parts of the song crystallize independently based, at least in part, on local error. This interpretation can explain the rapid changes in the gain of variability between syllables A and B documented in Chapter 2, as well as the lack of evidence that birds can easily increase the gain of local variability. This can be shown by an analogy:

Imagine walking through a corridor while looking outside via windows with variably sized openings: each time we cross a narrow window we are forced to look at the same image, but while crossing a wide-open window we have a range of images to choose from. Consequently, variability in the position of our eyes changes quickly as we walk by the windows, although the opening width of each window remains unchanged. By analogy, the rapid transitions we observed between variable and stereotyped song elements do not imply that the *range* of variability within each song element can be changed quickly, or reversed.

4.4.2 Hierarchical development

Taking articulatory dynamics into account, however, can complicate matters, since manipulating one vocal event could affect the acoustic states of neighboring events, and undo the learning. In this respect, learning temporary coarse song structure prior to the learning of temporary fine structure could potentially decrease such interactions and facilitate learning. As we have seen in Chapter 1, such transition from coarse segmentation of continuous action to fine segmentation may solve the problem of the trade-off between memory load (too fine segmentation) and the conflict between consolidation and exploration, which arises when segmentation is too coarse.

There is some evidence from previous studies that in articulation birds exposed to the normal song model (tutored birds) can learn more fine time structure than birds that have never heard a song model (untutored birds) (Méndez et al., 2010). Our results of exploratory variability analysis corroborate this view, since syllable types (clusters) became detectable in our data a few days prior to the appearance of robust intra-syllabic events. As we will see in the next chapter (Chapter 5) untutored birds do not show any signs of hierarchical development of their song (although the song does undergo some development and eventually consolidation).

4.4.3 Future directions

The results presented in this chapter are largely based on correlations between variability and either global or local error. The main result shows that correlations with the local error are stronger, supporting the hypothesis that errors from the song template are computed locally to gate vocal exploration dynamically in very short time scales. It would be interesting to further study the causality between vocal changes and exploration. Would exploratory variability increase when new sub-syllabic song elements are added to syllables? We have tried tutoring birds with sequential training analogous to the AAA → ABAB, in which a new sub-syllabic element was added to the “target”. However, we failed to induce intra syllabic changes in those birds. Possibly, the source syllable was too consolidated at the time when we altered the target syllable.. It is also possible that peripheral dynamics (e.g., co-articulation) imposed constraints were too strong to allow alteration of that syllable. However, this is not likely given the results by Tumer & Brainard (2007). A potentially better approach might be to set up an experiment analogous to Tumer and Brainard (Tumer and Brainard, 2007) where, as noted above, negative reinforcement was used to motivate birds to alter the fundamental frequency of a short song element. If such procedure is done in young birds, with plastic song, it is possible to test if the variability of that short song element will be higher than in the neighboring elements, when template match is not required. The alternative hypothesis is that the reason for our failure to induce intra syllable changes by altering the target is not due to motor constraints, but perhaps perceptual.